Be Responsible: A Novel Communications Scheme for Full-Duplex MIMO Radios

Diep N. Nguyen[1] and Marwan Krunz[2]
[1]Department of Engineering, Macquarie University, Australia, e-mail:diep.nguyen@mq.edu.au
[2]Department of Electrical and Computer Engineering, University of Arizona, USA, e-mail:krunz@email.arizona.edu

***Abstract*—Full-duplex (FD) radios have the potential to double a link's capacity. However, it has been recently reported that the network throughput gain of FD radios over half-duplex (HD) ones is unexpectedly marginal or even negative. This is because both ends of each link transmit at the same time, a set of concurrent FD links will experience more network interference (hence, reduction in the spatial reuse). This article identifies the unique advantages of FD radios and leverage multi-input multi-output (MIMO) communications to translate the FD spectral efficiency gain at the PHY level to throughput and power efficiency gain at the network layer. To that end, we first study the power minimization problem subject to rate demands in a FD-MIMO network. Sufficient conditions under which the FD network throughput can *asymptotically* double that of an HD network are then established. These conditions also guarantee the existence of a unique Nash Equilibrium that the game quickly converges to. By capturing "spatial signatures" of other radios, a FD-MIMO radio can instantly adjust its ongoing radiation pattern to avoid interfering with the reception directions at other radios. We exploit that to develop a novel MAC protocol that allows multiple FD links to concurrently communicate while adapting their radiation patterns to minimize network interference. The protocol does not require any feedback or coordination among nodes, but relies on the network interference perceived by these FD radios. Extensive simulations show that the proposed MAC design dramatically outperforms traditional FD-based CSMA protocols and HD radios w.r.t. both throughput and energy efficiency. A centralized algorithm for the FD network-wide transmit power minimization problem is also developed. Simulations show that, the proposed MAC protocol on average achieves almost the same power efficiency as the centralized algorithm. Interestingly, we even observe cases when the proposed distributed algorithm outperforms the centralized approach.**

***Index Terms*—Power efficiency, network throughput, full-duplex, MIMO, beamforming, Nash equilibrium, optimization, MAC.**

I. Introduction

Recent advances in self-interference suppression (SIS) allow a wireless device to transmit and receive simultaneously, i.e., perform full-duplex (FD) communications, on the same frequency [1] [2] and using the same antenna array [3] [4]. Over the last few years, various SIS techniques have been demonstrated, including antenna cancelation, analog RF, and digital cancelation (see [5] and references therein). Latest developments have successfully suppressed self-interference to the noise floor level for both single [3] and multi-antenna (i.e., MIMO) [4] devices. The spectral efficiency of an FD link has been shown to be nearly double that of a conventional half-duplex (HD) link [3] [4]. However, at the network layer, it has been reported that the network throughput gain under FD radios is unexpectedly marginal or even negative compared to HD-based systems [6] [7] [8]. This article attempts to translate the FD spectral efficiency gain at the link level into throughput and energy/power efficiency gains at the network layer.

Unlike HD radios, both ends of an FD link transmit at the same time. A set of mutually interfering FD links will now experience increased network interference and subsequently, severe reduction in spatial reuse. While previous works (e.g., [6] [7] [8]) identified the causes of throughput reduction in a network of FD radios, they failed to answer the question whether FD network throughput can be *ever* double that of HD. If it is possible, then under what conditions? Seeking an answer to this question is critical in designing efficient MAC protocols for FD-based multi-user systems.

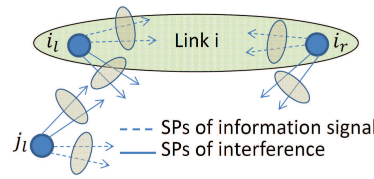


Fig. 1. Bidirectional link i is comprised of a left i_l and a right i_r radio. From the product of precoding and channel gain matrices from j_l to i_r (captured at RF chains of i_r) i_r can infer SPs of information signal of link j (implicitly embedded in precoding matrix of j_l) and SPs of interference induced on j_l from i_r (via channel reciprocity). i_l then can configure its radiation pattern to reduce as much interference as possible on j_r.

The FD capability not only improves the spectral efficiency but also allows a wireless device to *instantaneously* discern the medium while transmitting and instantly adjust its transmission strategy. This was leveraged to combat hidden/exposed terminals (in CSMA-based protocols) [9] or to improve the spectrum sensing/awareness in opportunistic access systems [10]. For MIMO communications, the *interference/signal perceived by a FD MIMO radio provides it with much detailed information (other than just the busy/idle status about the medium)*: It helps the node to partially infer "spatial signatures" (SPs) [11] [12] of the information signals intended to other nodes as well as the interference from the underlying node onto others. Specifically, the SP of the mth data (or interfering) stream is the mth column vector of the corresponding channel gain matrix, describing the spatial direction along which the stream's received power is maximized[1]. Note that in FD communications where the same antenna array is used to transmit and receive simultaneously [4], channel reciprocity holds (i.e., the channel gain matrix on one

This research was supported in part by the National Science Foundation (grants IIP-1432880, IIP-1265960, and CNS-1409172), the Army Research Office (grant W911NF-13-1-0302), Australian Research Council (Discovery Early Career Researcher Award), and the CSIRO Macquarie University Chair in Wireless Communications. This Chair has been established with funding provided by the Science and Industry Endowment Fund. Any opinions, findings, conclusions, or recommendations expressed in this paper are those of the author(s) and do not necessarily reflect the views of NSF or ARO.

[1]In MIMO communcations, signal alignment can be realized via *precoding*, in which the information symbol vector is pre-multiplied with a precoding matrix before being placed on a Tx antenna array. A precoding matrix, or precoder, is a matrix of complex elements whose phase and amplitude can be tuned to govern radiation directions/beams of the signal [11].

direction is the transpose of that of the other direction). Hence, as shown in Figure 1, the interference seen at the RF chains of radio i_r (product of interfering channel gain matrix from j_l to i_r and j_l's Tx precoder) contains SPs of information signal of link j (implicitly embedded in j_l's Tx precoder that is used to align j_l's signal along link j's SPs) and SPs of interference induced on j_l from i_r (via channel reciprocity).

Knowing the above SPs together with the ability to adjust the radiation pattern instantaneously (by tuning the phase and the amplitude of elements of its precoding matrix), a FD MIMO node can minimize its required Tx power and *"be responsible"* for reducing as much interference as possible on others. This ability does not exist in HD radios (where Tx and Rx have to take turn) or single-antenna radios, which cannot control their radiation beams. Note that existing works (e.g., [7] [8] [9]) considered a protocol model that does not allow links to coexist, and hence ignored the above advantage of FD MIMO radios which in fact facilitates the concurrent FD transmissions.

To establish the conditions that guarantee the superiority of FD over HD radios in a network setting, we consider the transmit power minimization problem subject to rate constraints, instead of the throughput maximization in [6] [7] [8]. Considering the power minimization problem allows us to derive sufficient conditions under which a set of rate demands can be met. We can then identify sufficient conditions for the FD network throughput to *asymptotically* reach twice that of a comparable HD network. Note that due to interference, the network-wide transmit power minimization problem is nonconvex. Hence, even with the availability of global network information, solving such a problem is prohibitively expensive. Existing approaches [13] [12] that solve the power minimization problem subject to SINR requirements for HD radios are inapplicable as our problem involves coupling matrix operations.

We formulate a noncooperative game in which FD radios are players who aim to meet their rate demands by optimizing their precoders. Instead of simply minimizing transmit powers as in HD radios, a FD radio minimizes the sum of transmit powers on its antennas, weighted by the transpose of its interference covariance matrix (perceived locally by the radio). Following the approach in [14], using recession analysis [15] and the variational inequality theory [16], we provide sufficient conditions under which a Nash Equilibrium (NE) exists and a set of rates can be met. We also prove that the NE is unique. At this NE, if a network of $2N$ HD radios (N links) can achieve a total throughput of dN bps (i.e., d bps per link), then with FD capability, for the same network/channel realization, $2N$ FD radios can achieve $2N(d-1)$ bps (i.e., $(d-1)$ bps per direction of a bidirectional link).

Simulations results show that for a given set of rate demands the proposed approach is much more power-efficient than when HD radios are used or when FDs do not exploit SPs. The total network transmit power under our approach is way less than N times that of the CSMA-based approach (where only one link is allowed to use the medium at a time) while the network throughput is N times higher. We also observe that the game converges quickly to its NE, facilitating the design of a practical MAC protocol (called FD-MAC). As a performance benchmark, we develop a centralized algorithm for the FD network-wide power minimization problem using augmented Lagrangian method. Our major contributions are as follows:

- We establish sufficient conditions under which FD radios can double the network throughput.
- We identify and exploit the unique advantages of FD radios to allow for coexistence of multiple links. This is done by having MIMO FD radios instantly discern the medium at a finer level (i.e., spatial signatures of other radios) and instantaneously adjust/adapt their radiation beams.
- We design an efficient MAC protocol for a network of FD radios. The proposed FD-MAC protocol does not require any feedback from or coordination between links, as precoders are designed using only local information. Via simulations, FD-MAC is shown to achieve almost the same performance as its centralized version (which aims to minimize the total network Tx power). FD-MAC yields much higher energy efficiency and throughout gain than that of CSMA-based approach as well as HD radios.
- We prove that there exists a unique NE to which FD-MAC converges. Simulations show that that the proposed distributed algorithm converges to this NE after a few iterations.

We use $(.)^*$ to denote the conjugate of a matrix, $(.)^H$ for its Hermitian transpose, tr$(.)$ for its trace, $|.|$ for the determinant, and $(.)^T$ for the matrix transpose. diag$_m(.)$ indicates the diagonal element (m,m) of a matrix, and sum$(.)$ gives the summation of all elements of a vector. Matrices and vectors are bold-faced.

In Section II, we present the network model and problem formulation. Conditions for the existence and uniqueness of the NE and rate-demand satisfaction, optimal precoders, MAC protocol are presented in Section III. The centralized algorithm is developed in IV. Numerical results are discussed in Section V, followed by concluding remarks in Section VI.

II. Network Model

Consider an ad hoc network of N FD-MIMO links. Both ends of each bidirectional link i operate simultaneously as transmitter and receiver. We differentiate the two FD radios/nodes of link i by denoting the left radio i_l and the right radio i_r. Without loss of generality, each node is equipped with M antennas (our analysis and results are applicable to the case where nodes have different numbers of antennas). The latest advances in SIS (e.g., [4]) allow the M-antenna array at each radio to transmit and receive simultaneously. Let $\mathbf{H}_{ii}^{rl}$ ($\mathbf{H}_{ii}^{lr}$) denote the $M \times M$ channel gain matrix of the left-to-right (right-to-left) direction of link i. Due to channel reciprocity, $\mathbf{H}_{ii}^{rl}$ is the transpose of $\mathbf{H}_{ii}^{lr}$. Each element of $\mathbf{H}_{ii}^{rl}$ is a multiplication of a distance- and frequency-dependent attenuation term and a random term that reflects multi-path fading (complex Gaussian variables with zero mean and unit variance). Let $\mathbf{H}_{ij}^{ll}$ and $\mathbf{H}_{ij}^{lr}$ denote the $M \times M$ interfering channel gain matrices from the radios j_l and j_r of link j on radio i_l of link i, respectively.

Let $\mathbf{G}_i^l$ and $\mathbf{G}_i^r$ denote the transmit precoding matrices at the left i_l and right i_r radios of link i, respectively. Let $\mathbf{x}_i^r$ denote the vector of transmit information symbols being placed on the antennas of radio i_r (for the right-to-left direction of link i). The received signal vector $\mathbf{y}_i^l$ at the antennas of radio i_l is:

$$\mathbf{y}_i^l = \mathbf{H}_{ii}^{lr}\mathbf{G}_i^r\mathbf{x}_i^r + \sqrt{g_{sis}}\mathbf{H}_{ii}^{ll}\mathbf{G}_i^l\mathbf{x}_i^l + \sum_{j=1|j\neq i}^{N}\left(\mathbf{H}_{ij}^{ll}\mathbf{G}_j^l\mathbf{x}_j^l + \mathbf{H}_{ij}^{lr}\mathbf{G}_j^r\mathbf{x}_j^r\right) + \mathbf{N_o} \tag{1}$$

where the first term is the intended signal, the second term is the self-interference induced by transmit chains of radio i_l (with g_{sis} and $\mathbf{H}_{ii}^{ll}$ being the self-interference suppression level and self-interference matrix, respectively), the third and forth terms represent interference from the left and right radios of link j, and $\mathbf{N_o}$ is an $M \times 1$ complex Gaussian noise vector with identity covariance matrix $\mathbf{I}$, representing the noise floor. Let c_i^l (c_i^r) denote the throughput received at the i_l (i_r) radio of link i. Treating interference from others radios as color noise, we have:

$$\begin{aligned} c_i^l &= \log|\mathbf{I} + \mathbf{G}_i^{r\,H}\mathbf{H}_{ii}^{lr\,H}\mathbf{Q}_i^{l\,-1}\mathbf{H}_{ii}^{lr}\mathbf{G}_i^r| \\ c_i^r &= \log|\mathbf{I} + \mathbf{G}_i^{l\,H}\mathbf{H}_{ii}^{rl\,H}\mathbf{Q}_i^{r\,-1}\mathbf{H}_{ii}^{rl}\mathbf{G}_i^l| \end{aligned} \tag{2}$$

where $\mathbf{Q}_i^r$ ($\mathbf{Q}_i^l$) is the noise-plus-interference covariance matrix at radio i_r (i_l):

$$\begin{aligned} \mathbf{Q}_i^r =& \mathbf{I} + g_{sis}\mathbf{H}_{ii}^{rr}\mathbf{G}_i^r\mathbf{G}_i^{r\,H}\mathbf{H}_{ii}^{rr\,H} \\ &+ \sum_{j=1|j\neq i}^{N}\left(\mathbf{H}_{ij}^{rl}\mathbf{G}_j^l\mathbf{G}_j^{l\,H}\mathbf{H}_{ij}^{rl\,H} + \mathbf{H}_{ij}^{rr}\mathbf{G}_j^r\mathbf{G}_j^{r\,H}\mathbf{H}_{ij}^{rr\,H}\right). \end{aligned}$$

The network-wide power minimization problem subject to radios/nodes' rate demands d_i^l (to be received by i_l) and d_i^r (to be received by i_r) is stated as follows:

$$\begin{aligned} \underset{\{\mathbf{G_i^l}, \mathbf{G_i^r}, \forall i\}}{\text{minimize}} \quad & \sum_{i=1}^{N}\{\text{tr}(\mathbf{G}_i^l\mathbf{G}_i^{lH}) + \text{tr}(\mathbf{G}_i^r\mathbf{G}_i^{rH})\} \\ \text{s.t.} \quad & \text{C1: } d_i^l \leq c_i^l, \ \ \forall i \\ & \text{C2: } d_i^r \leq c_i^r, \ \ \forall i \end{aligned} \tag{3}$$

III. Noncooperative Game Formulation

A. Formulation

The network-wide power minimization problem (3) is not convex, and hence is computationally expensive to solve even in a centralized manner. Additionally, collecting network information for (3) often requires excessive overhead. Existing works on HD-based systems formulate strategic noncooperative games where the players are transmitting nodes. A transmit precoder $\tilde{\mathbf{G}}_i^r$ of radio i_r is found from:

$$\begin{aligned} \underset{\{\mathbf{G}_i^r\}}{\text{minimize}} \quad & \text{tr}(\mathbf{G}_i^r\mathbf{G}_i^{rH}) \\ \text{s.t.} \quad & d_i^l \leq c_i^l. \end{aligned} \tag{4}$$

similarly, the transmit precoder $\mathbf{G}_i^l$ of radio i_l can be found by solving:

$$\begin{aligned} \underset{\{\mathbf{G}_i^l\}}{\text{minimize}} \quad & \text{tr}(\mathbf{G}_i^l\mathbf{G}_i^{lH}) \\ \text{s.t.} \quad & d_i^r \leq c_i^r \end{aligned}$$

As aforementioned, a FD radio, i_r, can use its receive chains to gauge how much interference its antennas induce on others. Specifically, consider the transpose of the covariance matrix of interference-plus-noise perceived by radio i_r:

$$\begin{aligned} \mathbf{Q}_i^{rT} =& \mathbf{I} + g_{sis}(\mathbf{H}_{ii}^{rr}\mathbf{G}_i^r\mathbf{G}_i^{r\,H}\mathbf{H}_{ii}^{rr\,H})^T \\ &+ \sum_{j=1|j\neq i}^{N}\left((\mathbf{G}_j^{l\,T}\mathbf{H}_{ji}^{lr})^H(\mathbf{G}_j^{l\,T}\mathbf{H}_{ji}^{lr}) + (\mathbf{G}_j^{rT}\mathbf{H}_{ji}^{rr})^H(\mathbf{G}_j^{rT}\mathbf{H}_{ji}^{rr})\right) \end{aligned}$$

Let:

$$\begin{aligned} \mathbf{S}_i^r \stackrel{\text{def}}{=}& g_{sis}(\mathbf{H}_{ii}^{rr}\mathbf{G}_i^r\mathbf{G}_i^{r\,H}\mathbf{H}_{ii}^{rr\,H})^T \\ &+ \sum_{j=1|j\neq i}^{N}\left((\mathbf{G}_j^{l\,T}\mathbf{H}_{ji}^{lr})^H(\mathbf{G}_j^{l\,T}\mathbf{H}_{ji}^{lr}) + (\mathbf{G}_j^{rT}\mathbf{H}_{ji}^{rr})^H(\mathbf{G}_j^{rT}\mathbf{H}_{ji}^{rr})\right). \end{aligned}$$

$\mathbf{S}_i^r$ contains *spatial signatures* of the interference signal induced by radio i_r onto radio j_l ($\mathbf{H}_{ji}^{lr}$) and radio j_r ($\mathbf{H}_{ji}^{rr}$). $\mathbf{S}_i^r$ also captures the SPs of information signal ($\mathbf{H}_{jj}^{lr}$, $\mathbf{H}_{jj}^{rl}$) intended for radios j_l and j_r that are implicitly embedded in transmit precoders $\mathbf{G}_j^l$ and $\mathbf{G}_j^r$ (as radio j_l aligns its data streams with the sub-channels directions of $\mathbf{H}_{jj}^{rl}$ while $\mathbf{H}_{jj}^{lr} = \mathbf{H}_{jj}^{rl\,T}$).

Intuitively, for an interfering channel, SPs capture the *vulnerable* directions that interference is most harmful. For an information signal, SPs are directions along which the transmit/receive beamformers should align the signal to maximize the signal's received power [11][2]. Exploiting knowledge of other nodes' SPs that is learned while transmitting, an FD radio can meet its rate demand while minimizing both transmit power and interference induced on other radios. To that end, the precoder of radio i_r can be obtained from:

$$\begin{aligned} \underset{\{\mathbf{G}_i^r\}}{\text{minimize}} \quad & \text{tr}(\mathbf{G}_i^r\mathbf{G}_i^{rH}) + \text{tr}(\mathbf{G}_i^r\mathbf{S}_i^r\mathbf{G}_i^{rH}) \\ \text{s.t.} \quad & d_i^l \leq c_i^l \end{aligned} \tag{5}$$

where $\text{tr}(\mathbf{G}_i^r\mathbf{G}_i^{rH}) + \text{tr}(\mathbf{G}_i^r\mathbf{S}_i^r\mathbf{G}_i^{rH}) = \text{tr}(\mathbf{G}_i^r\mathbf{Q}_i^{rT}\mathbf{G}_i^{rH})$ is interpreted as the summation of transmit power and interference caused by i_r.

B. Nash Equilibrium Existence and Uniqueness

The two games (4) and (5) have identical strategic space, defined as the union of all players' strategic spaces [17], which is shaped by rate constraints C1 and C2. We thus can focus on analyzing the game (5), game (4) then follows by replacing $\mathbf{Q}_i^{rT}$ in (5) with the identity matrix $\mathbf{I}$.

Optimizing the precoder $\mathbf{G}_i^r$ of radio i_r embodies computing the optimal radiation directions and power allocation across i_r's antennas. We can rewrite $\mathbf{G}_i^r$ as:

$$\mathbf{G}_i^r = \tilde{\mathbf{G}}_i^r \times \mathbf{P}_i^{r\,1/2} \tag{10}$$

where $\tilde{\mathbf{G}}_i^r$ is an $M \times M$ unit-norm column matrix, controlling radiation directions of radio i_r. $\mathbf{P}_i^r$ is an $M \times M$ diagonal matrix whose diagonal element $\mathbf{P}_i^r(m, m)$ is the power allocated on mth data stream of radio i_r.

Let $\boldsymbol{p}_i^r \stackrel{\text{def}}{=} [\mathbf{P}_i^r(1,1), \mathbf{P}_i^r(2,2), \ldots, \mathbf{P}_i^r(M,M)]$ denote the power allocation of radio i_r for its various data streams. Let $\boldsymbol{p} \stackrel{\text{def}}{=} [\boldsymbol{p}_1^l, \boldsymbol{p}_1^r, \ldots, \boldsymbol{p}_N^l, \boldsymbol{p}_N^r] \in \mathbb{R}_+^{2NM}$ denote the power allocation on all data streams of all ($2N$) radios in the network. $\mathbf{Q}_i^r$ is positive definite, and so is its transpose. The objective function in (5) is non-decreasing in every element of $\boldsymbol{p}_i^r$. Thus, at a NE of the game (if one exists), the rate demand inequality constraint becomes active (i.e., turns to equality). If not, the radio is able to lower its transmit power and reduce the objective function in (5) while fulfilling its rate demand. This defines a feasible set for $\boldsymbol{p}$, denoted by $\mathbb{P}_{feasible}(\boldsymbol{d})$ in (6), corresponding to a given requested rate profile $\boldsymbol{d} \stackrel{\text{def}}{=} [d_1^r, d_1^l, \ldots, d_N^r, d_N^l]$ at a NE. We first prove that if a rate demand profile can be supported with a finite power vector $\boldsymbol{p}$, then the game (5) admits at least one NE.

Theorem 1: If all rate demands can be supported, then the game (5) admits at least one NE.

Proof: If all rate demands can be met, the strategic space of game (5) must be nonempty. Additionally, it can be verified that the strategic space of each player (defined by the rate constraint) in (5) is convex, as the achievable throughput is concave w.r.t. a radio's precoder. Since rate demands can be

[2]For example, the minimum mean square error (MMSE) receiver at j_l is capacity-optimal by setting its receive beamformer to $\mathbf{Q}_j^{l\,-1}\mathbf{H}_{jj}^{lr}$ [11]

supported with finite transmit powers, we can add technical constraints on radios' power budget. This makes these strategic spaces compact. In short, the strategic space of (5) is nonempty, convex, and compact. Moreover, the player's payoff in the objective function of (5) is convex. Hence, (5) admits at least one NE [17]. □

For the existence of a NE, it suffices to find conditions under which the feasible set $\mathbb{P}_{feasible}(\boldsymbol{d})$ of $\boldsymbol{p}$ is nonempty and bounded. This is formally stated in the following theorem:

Theorem 2: Let $\boldsymbol{\Gamma}$ be a $2N \times 2N$ matrix that is defined in (7). If $\boldsymbol{\Gamma}$ is a P-matrix[3], then $\mathbb{P}_{feasible}(\boldsymbol{d}) \in \mathbb{R}_{+}^{2NM}$ is nonempty and bounded, and hence the game (5) admits at least one NE.

Proof: We first prove that $\mathbb{P}_{feasible}(\boldsymbol{d})$ contains at least one bounded vector $\boldsymbol{p} \in \mathbb{R}_{+}^{NKM}$ or the the rate remand can be met with finite transmit power.

Lemma 1: Given that $\boldsymbol{\Gamma}$ is a P-matrix, there exists at least one bounded vector $\boldsymbol{p} \in \mathbb{P}_{feasible}(\boldsymbol{d}) \in \mathbb{R}_{+}^{2NM}$.

Proof: See Appendix A. □

Next, to show that $\mathbb{P}_{feasible}(\boldsymbol{d})$ is bounded, we rely on the concept of an asymptotic cone of a nonempty set in recession analysis [15]. Specifically, for a nonempty set $\mathbb{P} \in \mathbb{R}_{+}^{N}$, its *asymptotic cone*, $\mathbb{P}_{asymp}$, is comprised of vectors $\boldsymbol{f} \in \mathbb{R}_{+}^{N}$, called limit directions. Each limit direction vector $\boldsymbol{f}$ is defined through the existence of a sequence of vectors $\boldsymbol{p}_n \in \mathbb{P}$ and a sequence of scalars ν_n tending to $+\infty$ such that [15]:

$$\lim_{n\to\infty} \frac{\boldsymbol{p}_n}{\nu_n} = \boldsymbol{f}. \tag{11}$$

$\mathbb{P}$ is bounded if its asymptotic cone $\mathbb{P}_{asymp} = \{\mathbf{0}\}$ [15]. To show that $\mathbb{P}_{feasible}(\boldsymbol{d})$ is bounded, it suffices to prove that its asymptotic cone $\mathbb{P}_{asymp}(\boldsymbol{d})$ contains only the zero vector. The asymptotic cone $\mathbb{P}_{asymp}(\boldsymbol{d})$ is formally defined in (8).

Since $\mathbb{P}_{feasible}(\boldsymbol{d})$ has at least one bounded $\boldsymbol{p}$ (Lemma 1), by the definition of limit directions, the vector zero $\mathbf{0}$ belongs to its asymptotic cone $\mathbb{P}_{asymp}(\boldsymbol{d})$. We now construct a set $\mathbb{P}(\boldsymbol{d})$ of which $\mathbb{P}_{asymp}(\boldsymbol{d})$ is a subset and prove that $\mathbb{P}(\boldsymbol{d}) = \{\mathbf{0}\}$ if $\boldsymbol{\Gamma}$ is a P-matrix.

Lemma 2: If $\boldsymbol{f} \in \mathbb{P}_{asymp}(\boldsymbol{d})$ then $\boldsymbol{f}$ belongs to $\mathbb{P}(\boldsymbol{d})$, defined in (9).

Proof: See Appendix B. □

[3]A P-matrix is one of which all principal minors are positive [18].

Assuming that there exists at least one $\boldsymbol{f} \neq \mathbf{0}$ and that $\boldsymbol{f} \in \mathbb{P}(\boldsymbol{d})$, we have:

$$\boldsymbol{\Gamma} \times [\text{tr}(\mathbf{G}_1^{lH}\mathbf{G}_1^{l}), \ldots, \text{tr}(\mathbf{G}_N^{rH}\mathbf{G}_N^{r})]^T \leq \mathbf{0}. \tag{12}$$

As $\boldsymbol{\Gamma}$ is a P-matrix and $[\text{tr}(\mathbf{G}_1^{lH}\mathbf{G}_1^{l}), \ldots, \text{tr}(\mathbf{G}_N^{rH}\mathbf{G}_N^{r})]^T$ is a nonnegative vector, (12) implies that $\text{tr}(\mathbf{G}_i^{lH}\mathbf{G}_i^{l}) = 0$ and $\text{tr}(\mathbf{G}_i^{rH}\mathbf{G}_i^{r}) = 0\ \forall i$ [18] or $\boldsymbol{f} = \mathbf{0}$. This contradicts the above assumption. Hence, $\mathbb{P}(\boldsymbol{d})$ and its subset $\mathbb{P}_{asymp}(\boldsymbol{d})$ equal to $\{\mathbf{0}\}$. Theorem 2 is proved. □

Intuitions behind Theorem 2 can be drawn as follows. If the diagonal elements of $\boldsymbol{\Gamma}$ are positive, then a sufficient condition for $\boldsymbol{\Gamma}$ to be a P-matrix is $|\boldsymbol{\Gamma}(i,i)| \geq \sum_{j\neq i} |\boldsymbol{\Gamma}(i,j)|$ (i.e., row diagonally dominant) [18]. Hence, the following inequality guarantees that game (5) has at least one NE:

$$\frac{M\det(\mathbf{H}_{ii}^{lrH}\mathbf{H}_{ii}^{lr})^{\frac{1}{M}}}{g_{sis}\text{tr}(\mathbf{H}_{ii}^{llH}\mathbf{H}_{ii}^{ll}) + \sum_{j=1|j\neq i}^{N} (\text{tr}(\mathbf{H}_{ij}^{llH}\mathbf{H}_{ij}^{ll}) + \text{tr}(\mathbf{H}_{ij}^{lrH}\mathbf{H}_{ij}^{lr}))} \geq (2^{d_i^l}-1), \forall \text{ i}. \tag{13}$$

To better interpret inequality (13), let's assume SIS is perfect (i.e., g_{sis} is sufficiently small to be neglected) and rewrite $\mathbf{H}_{ii}^{lr} = \frac{1}{\sqrt{s_{ii}^{lr\,n}}}\bar{\mathbf{H}}_{ii}^{lr}$ where n is the path loss exponent, ${s_{ii}}^{lr}$ is the transmission distance from radio i_r to radio i_l, and $\bar{\mathbf{H}}_{ii}^{lr}$ is a complex Gaussian matrix with with zero mean and unit variance. Inequality (13) can be rewritten as:

$$\frac{M\det(\bar{\mathbf{H}}_{ii}^{lrH}\bar{\mathbf{H}}_{ii}^{lr})^{\frac{1}{M}}}{\sum_{j=1|j\neq i}^{N} (\frac{s_{ii}^{lr\,n}}{s_{ij}^{ll\,n}}\text{tr}(\bar{\mathbf{H}}_{ij}^{llH}\bar{\mathbf{H}}_{ij}^{ll}) + \frac{s_{ii}^{lr\,n}}{s_{ij}^{lr\,n}}\text{tr}(\bar{\mathbf{H}}_{ij}^{lrH}\bar{\mathbf{H}}_{ij}^{lr}))} \geq (2^{{d_i}^l}-1)\ \forall\ i. \tag{14}$$

The nominator of the LHS in (14) represents the strength of the channel gain matrix from i_r to i_l, while its denominator describes the strength of (interfering) channel gain matrices from all other radios j_l and j_r ($j \neq i$) on radio i_l. For the game (5) to have at least one NE, the multi-user interference should not be too strong. This is the case if the (transmission) distance s_{ii}^{lr} between the i_l and i_r is small enough compared with (interfering) distances (s_{ij}^{ll} and s_{ij}^{lr}) between i_l and other radios (other than i_r), the channel gain matrix of link i is full-

$$\mathbb{P}_{feasible}(\boldsymbol{d}) \stackrel{\text{def}}{=} \left\{\boldsymbol{p} \in \mathbb{R}_{+}^{2NM} | c_i^l(\boldsymbol{p}) \stackrel{\text{def}}{=} \log|\mathbf{I} + \mathbf{G}_i^{rH}\mathbf{H}_{ii}^{lrH}{\mathbf{Q}_i^l}^{-1}\mathbf{H}_{ii}^{lr}\mathbf{G}_i^{r}| = d_i^l, \forall i_r, \quad , i_l\right\} \tag{6}$$

$$\boldsymbol{\Gamma} \stackrel{\text{def}}{=} \begin{bmatrix} |\mathbf{H}_{11}^{lrH}\mathbf{H}_{11}^{lr}|^{\frac{1}{M}} & -g_{sis}(2^{d_1^l}-1)\frac{\text{tr}(\mathbf{H}_{11}^{llH}\mathbf{H}_{11}^{ll})}{M} & -(2^{d_1^l}-1)\frac{\text{tr}(\mathbf{H}_{12}^{llH}\mathbf{H}_{12}^{ll})}{M} & \cdots & -(2^{d_1^l}-1)\frac{\text{tr}(\mathbf{H}_{1N}^{llH}\mathbf{H}_{1N}^{ll})}{M} & -(2^{d_1^l}-1)\frac{\text{tr}(\mathbf{H}_{1N}^{lrH}\mathbf{H}_{1N}^{lr})}{M} \\ -g_{sis}(2^{d_1^r}-1)\frac{\text{tr}(\mathbf{H}_{11}^{rrH}\mathbf{H}_{11}^{rr})}{M} & |\mathbf{H}_{11}^{rlH}\mathbf{H}_{11}^{rl}|^{\frac{1}{M}} & -(2^{d_1^r}-1)\frac{\text{tr}(\mathbf{H}_{12}^{rlH}\mathbf{H}_{12}^{rl})}{M} & \cdots & -(2^{d_1^r}-1)\frac{\text{tr}(\mathbf{H}_{1N}^{rlH}\mathbf{H}_{1N}^{rl})}{M} & -(2^{d_1^r}-1)\frac{\text{tr}(\mathbf{H}_{1N}^{rrH}\mathbf{H}_{1N}^{rr})}{M} \\ \vdots & \vdots & \ddots & \vdots & & \\ -(2^{d_N^r}-1)\frac{\text{tr}(\mathbf{H}_{N1}^{rlH}\mathbf{H}_{N1}^{rl})}{M} & -(2^{d_N^r}-1)\frac{\text{tr}(\mathbf{H}_{N1}^{rrH}\mathbf{H}_{N1}^{rr})}{M} & -(2^{d_N^r}-1)\frac{\text{tr}(\mathbf{H}_{N2}^{rlH}\mathbf{H}_{N2}^{rl})}{M} & \cdots & -g_{sis}(2^{d_N^r}-1)\frac{\text{tr}(\mathbf{H}_{NN}^{rrH}\mathbf{H}_{NN}^{rr})}{M} & |\mathbf{H}_{NN}^{rlH}\mathbf{H}_{NN}^{rl}|^{\frac{1}{M}} \end{bmatrix} \tag{7}$$

$$\mathbb{P}_{asymp}(\boldsymbol{d}) \stackrel{\text{def}}{=} \left\{\boldsymbol{f} \in \mathbb{R}_{+}^{2NM} | \exists\{\boldsymbol{p}_n\} \in \mathbb{P}_{feasible}(\boldsymbol{d}) \text{ and } \{\nu_n\} \to +\infty \text{ so that } \lim_{n\to\infty}\frac{\boldsymbol{p}_n}{\nu_n} = \boldsymbol{f}\right\} \tag{8}$$

$$\mathbb{P}(\boldsymbol{d}) \stackrel{\text{def}}{=} \left\{\boldsymbol{f} \in \mathbb{R}_{+}^{2NM} | {c'}_i^l(\boldsymbol{f}) \stackrel{\text{def}}{=} \log\left(1 + \frac{\text{tr}(\mathbf{G}_i^{rH}\mathbf{G}_i^{r})|\mathbf{H}_{ii}^{lrH}\mathbf{H}_{ii}^{lr}|^{\frac{1}{M}}}{g_{sis}\frac{\text{tr}(\mathbf{H}_{ii}^{llH}\mathbf{H}_{ii}^{ll})}{M}\text{tr}(\mathbf{G}_i^{lH}\mathbf{G}_i^{l}) + \sum_{j=1|j\neq i}^{N}(\frac{\text{tr}(\mathbf{H}_{ij}^{llH}\mathbf{H}_{ij}^{ll})}{M}\text{tr}(\mathbf{G}_j^{lH}\mathbf{G}_j^{l}) + \frac{\text{tr}(\mathbf{H}_{ij}^{lrH}\mathbf{H}_{ij}^{lr})}{M}\text{tr}(\mathbf{G}_j^{rH}\mathbf{G}_j^{r}))}\right) \leq d_i^l, \forall i\right\} \tag{9}$$

rank (this is often the case in a rich-scattering environment) and its requested rate is not too high. The acceptable multi-user interference is explicitly quantified in (13), and is a function of the rate demand d_i^l of radio i_l. For higher rate demands, inequality (13) becomes stringent, meaning that network interference must be lowered.

Remark 1: The denominator of (14) captures the interference channel gains from radios on both left j_l and right j_r sides of all links $j \neq i$ to i_l. If all FD radios choose to operate in HD mode, e.g., all left radios Tx and all right radios Rx, then the second term in the denominator should disappear. Consequently, the denominator reduces roughly by one half. If (14) holds for the new denominator for all radios with rate demands $d_i^r = d$ and $d_i^l = 0$ (as all left radios Tx) for all i, then when all radios return to FD mode, (14) should also hold for all radios with rate demands $d_i^r = d - 1$ and $d_i^r = d - 1$. This is because $\frac{(2^d - 1)}{2} > 2^{d-1} - 1$. The network throughput in the FD case is then $2N(d-1)$, which is asymptotically twice that of the HD case (Nd). Hence, we can conjecture that if (14) holds for all radios, network throughput can double with FD radios. Conditions in (14) are also in line with the findings in [7] [8] where the authors observed that FD radios outperforms HD ones (in terms of network throughput) if network interference is mild (i.e., interfering links are sufficiently separated from each other). However, [7] [8] did not quantify how mild network interference should be for FD radios to double the network throughput.

To analyze the uniqueness of the NE, we rely on variational inequalities (VI) theory, casting (5) as a VI problem. A tutorial on VI can be found in [16] and references therein.

Theorem 3: If game (5) has a NE, this NE is unique.

Proof: We prove that the mapping of the equivalent VI problem of (5) is continuous uniformly-P function. Hence, if a NE exists, it is unique. See Appendix C for details. □

Remark 2: The conditions in Theorem 2 or inequality (14) are sufficient but not necessary. FD radios can still double network throughput even when these conditions do not hold. Theorem 3 indicates that (5) does not have multiple NEs. In simulations, the game always converges to its unique NE as long as the rate demands are not unreasonably high.

C. Best Response

The optimal precoder $\mathbf{G}_i^r$ of radio i_r is obtained by solving (5). Notice that (5) is convex, hence can be solved efficiently with the interior-point method. To gain insights into how power is allocated over i_r's antennas, we follow the approach in [19] by using Hadamard inequality [20]. Specifically, the Lagrange function of (5) is:

$$\begin{aligned} & L_i^r(\tilde{\mathbf{G}}_i^r, \gamma_i^l) && (15a)\\ &= \mathrm{tr}(\mathbf{G}_i^r \mathbf{Q}_i^{rT} \mathbf{G}_i^{rH}) + \gamma_i^l (d_i^l - \log |\mathbf{I} + \mathbf{G}_i^{rH} \mathbf{H}_{ii}^{lrH} \mathbf{Q}_i^{l^{-1}} \mathbf{H}_{ii}^{lr} \mathbf{G}_i^r|) \\ &= \mathrm{tr}(\bar{\mathbf{G}}_i^r \bar{\mathbf{G}}_i^{rH}) + \gamma_i^l (d_i^l - \log |\mathbf{I} + \bar{\mathbf{G}}_i^{rH} \mathbf{E}_i^{r^{-1}} \mathbf{H}_{ii}^{lrH} \mathbf{Q}_i^{l^{-1}} \mathbf{H}_{ii}^{lr} \mathbf{E}_i^{rH^{-1}} \bar{\mathbf{G}}_i^r|) \\ &\leq \mathrm{tr}(\bar{\mathbf{G}}_i^r \bar{\mathbf{G}}_i^{rH}) && (15b)\\ &\quad + \gamma_i^l (d_i^l - \sum_{m=1}^{M} \log(1 + \mathrm{diag}_m(\bar{\mathbf{G}}_i^{rH} \mathbf{E}_i^{r^{-1}} \mathbf{H}_{ii}^{lrH} \mathbf{Q}_i^{l^{-1}} \mathbf{H}_{ii}^{lr} \mathbf{E}_i^{rH^{-1}} \bar{\mathbf{G}}_i^r|))) \end{aligned}$$

where γ_i^l, γ_i^r are nonnegative Lagrangian multipliers and $\bar{\mathbf{G}}_i^{rH} = \mathbf{G}_i^{rH} \mathbf{E}_i^r$ with Cholesky decomposition $\mathbf{Q}_i^{rT} = \mathbf{E}_i^r \mathbf{E}_i^{rH}$. The last inequality is obtained by applying Hadamard inequality [20].

Problem (5) can be solved by finding the maximum of its lower bound $L_i^r(\tilde{\mathbf{G}}_i^r, \gamma_i^l)$. Inequality (15b) becomes an equality if there exists an orthonormal matrix $\bar{\mathbf{G}}_i^r$ that diagonalizes $\mathbf{E}_i^{r^{-1}} \mathbf{H}_{ii}^{lrH} \mathbf{Q}_i^{l^{-1}} \mathbf{H}_{ii}^{lr} \mathbf{E}_i^{rH^{-1}}$. After a few manipulations, we can prove that the optimal $\mathbf{G}_i^r$ must be in the form a generalized eigen matrix of $\mathbf{H}_{ii}^{lrH} \mathbf{Q}_i^{l^{-1}} \mathbf{H}_{ii}^{lr}$ and $\mathbf{Q}_i^{rT}$. This is realized by setting $\tilde{\mathbf{G}}_i^r$ in (10) as a unit-norm generalized eigen matrix of $\mathbf{H}_{ii}^{lrH} \mathbf{Q}_i^{l^{-1}} \mathbf{H}_{ii}^{lr}$ and $\mathbf{Q}_i^{rT}$. It follows from [21] that:

$$\begin{aligned} &\tilde{\mathbf{G}}_i^{rH} \mathbf{H}_{ii}^{lrH} \mathbf{Q}_i^{l^{-1}} \mathbf{H}_{ii}^{lr} \tilde{\mathbf{G}}_i^r = \mathbf{\Pi}_i^l \text{ and} \\ &\tilde{\mathbf{G}}_i^{rH} \mathbf{Q}_i^{rT} \tilde{\mathbf{G}}_i^r = \mathbf{\Omega}_i^r \end{aligned} \tag{16}$$

where $\mathbf{\Pi}_i^l$ and $\mathbf{\Omega}_i^r$ are diagonal matrices.

The Lagrangian $L_i^r(\mathbf{G}_i^r, \gamma_i^l)$ becomes:

$$\begin{aligned} &L_i^r(\mathbf{G}_i^r, \gamma_i^l) = \\ &\sum_{m=1}^{M} (\mathrm{diag}_m(\mathbf{\Omega}_i^r) \mathbf{P}_i^r(m,m) - \gamma_i^l \log(1 + \mathbf{P}_i^r(m,m) \mathrm{diag}_m(\mathbf{\Pi}_i^l))). \end{aligned}$$

The optimal power for data stream m is obtained by equating the derivative of $L_i^r(\mathbf{G}_i^r, \gamma_i^l)$ to zero. Accordingly:

$$\mathbf{P}_i^r(m,m) = \max\left(0, \gamma_i^l \frac{1}{\mathrm{diag}_m(\mathbf{\Omega}_i^r)} - \frac{1}{\mathrm{diag}_m(\mathbf{\Pi}_i^r)}\right) \tag{17}$$

where the Lagrange multiplier γ_i^l is found (e.g., using bisection search) to meet the rate demand d_i^l.

From (17), more power is allocated on data streams with lower $\mathrm{diag}_m(\mathbf{\Omega}_i^r)$ and higher $\mathrm{diag}_m(\mathbf{\Pi}_i^r)$. This means more power is allocated to higher-gain streams and less power on directions that cause higher interference to others.

D. MAC Protocol

We briefly present a MAC protocol, called FD-MAC, that implements the game (5) in a distributed fashion. Unlike typical CSMA-based protocols, FD-MAC exploits information perceived by FD radios to enable concurrent transmissions on multiple links. Each transmission session in FD-MAC consists of two phases: a training phase and a data transmission phase. In the first phase, an FD radio A with packets to send transmits a hand-shaking message (HSK), containing a training sequence for CSI estimation purposes to rendezvous with its intended radio B. HSKs are sent at the lowest rate, referred to as a base rate, so as to improve the chance it is successfully decoded.

As each FD radio can transmit and receive at the same time and a two-way channel exists between the two radios of a bidirectional link, FD radios of a link can instantaneously update each other regarding CSI as well as noise-plus-interference covariance matrix. This information is needed to solve (5). If either A or B fails to transmit or receive at the base rate to hand-shake with its intended partner, it then holds off for a random duration before trying again. Upon receiving an HSK, radio B replies with a message to trigger the training process by solving problem (5) to achieve the rate demand. The data transmission phase ensues with multiple packets. Note that under FD-MAC, radios of different links do not need to coordinate or exchange any signalling packets, and precoders are computed using only local information.

To ensure that the training phase in FD-MAC is not too long, the iterating process should converge after a reasonable time. Although we cannot prove the convergence of the game (5) under arbitrary/asynchronous updates, simulations show that the game converges even if some radios sporadically skip updating their precoders. The following theorem claims the convergence of the game to its unique NE under sequential (Gauss-Seidel) iterations.

Theorem 4: Under the sequential (Gauss-Seidel) iterations, the game (5) converges to its NE.

Proof: We can follow the routine in [22] [23] to construct a Lyapunov function of the precoding matrices and show that this function is non-increasing and lower-bounded. The convergence point must be a NE; otherwise one user can still unilaterally reduce its transmit power and that violates the convexity of (5). □

IV. Network-wide Problem

To seek a performance benchmark, in this section, we use the augmented Lagrange multiplier method [24] to derive the centralized algorithm for the network-wide problem (3). The augmented Lagrange of (3) is given in (20), where $q_i^l \overset{\text{def}}{=} d_i^l - c_i^l$, $q_i^r \overset{\text{def}}{=} d_i^r - c_i^r$. p is a positive penalty factor for violating rate constraints. At an optimal solution, (21) holds.

Since q_j^l is continuously differentiable w.r.t every entry of $\tilde{\mathbf{G}}_j^r$, the third and fourth terms in (21) are also continuously differentiable [24]. Their derivatives are as follows:

$$\frac{\partial\{(\max\{0,\gamma_i^l+pq_i^l\})^2\}}{\partial \mathbf{G}_j^{r*}}=\begin{cases}0 & \text{if } \gamma_i^l+pq_i^l \le 0\\ -2p\frac{\partial c_i^l}{\partial \mathbf{G}_i^{r*}} & \text{if } \gamma_i^l+pq_i^l > 0 \text{ and } i=j\\ -2p\frac{\partial c_i^l}{\partial \mathbf{G}_j^{r*}} & \text{if } \gamma_i^l+pq_i^l > 0 \text{ and } i\neq j\end{cases}$$

$$\frac{\partial\{(\max\{0,\gamma_i^r+pq_i^r\})^2\}}{\partial \mathbf{G}_j^{r*}}=\begin{cases}0 & \text{if } \gamma_i^r+pq_i^r \le 0 \text{ or } j=i\\ -2p\frac{\partial c_i^r}{\partial \tilde{\mathbf{G}}_j^{r*}} & \text{if } \gamma_i^r+pq_i^r > 0 \text{ and } j\neq i\end{cases}$$

where:

$$\frac{\partial c_i^l}{\partial \mathbf{G}_i^{r*}} = -\mathbf{H}_{ii}^{lrH}(\mathbf{Q}_i^l + \mathbf{H}_{ii}^{lr}\mathbf{G}_i^r\mathbf{G}_i^{rH}\mathbf{H}_{ii}^{lrH})^{-1}\mathbf{H}_{ii}^{lr}\mathbf{G}_i^r \quad (18)$$

and

$$\frac{\partial c_i^l}{\partial \mathbf{G}_j^{r*}} = \mathbf{H}_{ij}^{lrH}\mathbf{Q}_i^{l^{-1}}\mathbf{H}_{ij}^{lr}[(\mathbf{G}_i^r\mathbf{G}_i^{rH})^{-1} + \mathbf{H}_{ii}^{lrH}\mathbf{Q}_i^{l^{-1}}\mathbf{H}_{ii}^{lr}]^{-1}\mathbf{H}_{ii}^{lrH}\mathbf{Q}_i^{l^{-1}}\mathbf{H}_{ij}^{lr}\mathbf{G}_j^r \quad (19)$$

We use the gradient search algorithm with Armijo step size [24] to find $(\mathbf{G}_i^l, \mathbf{G}_i^r, \gamma_i^l, \gamma_i^r, p)$ such that (21) holds for all radios. The running time can be high as it involves NM^2 complex variables (or $2NM^2$ real ones).

V. Simulations Results

We numerically evaluate the performance of the above centralized and distributed algorithms using MATLAB simulations. We compare total network power required to meet a given set of rate demands of the centralized algorithm with that of distributed algorithms when FD-capability in capturing spatial signatures is exploited (game (5), FD-MAC) and not exploited (game (4), namely "without SPE"). We also compare total required power under the proposed FD-MAC protocol with that when FD links take turn to access the channel (i.e., CSMA-based protocols). 8 pairs of radios are randomly placed in a field of $500 \times 500\ m^2$. Each radio has 4 antennas. Channel bandwidth is 20 MHz. Noise floor is set as -90dBm/Hz. The channel fading is flat with a free-space attenuation factor of 2. All algorithms have identical initializations of precoding matrices.

Figure 2 shows snapshots of radiation patterns of FD radios under different algorithms (at their converged points). As can be seen, when radios cooperate so that (3) can be solved in a centralized manner, we visually notice that FD radios try to steer their beams away from other unintended ones. This is also observed when SPs are exploited in the distributed FD-MAC algorithm when nodes minimize their Tx power weighted by SPs of others. Interestingly, the beam patterns of the distributed FD-MAC are quite similar to that of the centralized algorithm, suggesting the efficiency of using SPs. These two algorithms seem induce less network interference, compared with beam patterns of the case when SPs are not exploited (game (4)).

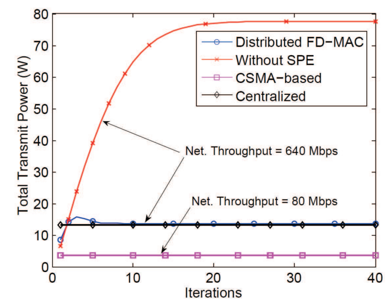


Fig. 3. Total network transmit power under different algorithms for rate demand of 40 Mbps per radio.

Figure 3 depicts total network transmit power required when all radios demand a rate of 40 Mbps (i.e., 2 bps/Hz). Notice that the CSMA-based approach where FD links take turn to access the medium (i.e., only one FD link operates at a time) requires the least Tx power. This is because in such cases, a link does not need to cope with interference from others, but at the expense of the lowest network throughput (equals to that of one FD link's, 80 Mbps). The proposed FD-MAC algorithm converges after about 9 iterations and consumes almost the same (barely higher than) as total power under the centralized algorithm. This seems to agree with the radiation behavior observed in Figure 2. Compared with the CSMA-based approach, by advocating concurrent links' transmission, FD-MAC requires about 5 times higher Tx power but attains 640 Mbps network throughput (8 times higher). This gain is very significant due to the fact that throughput/rate does not scale w.r.t. transmit power.

FD-MAC outperforms CSMA approach as a link does not give up when the medium is busy. Instead, it proceeds but in a responsible/careful way by exploiting SPs to minimize interference to ongoing transmissions. We also observe that for the same amount throughout 640 Mbps, if SPs are not exploited, the required transmit power is 77.4 W (compared to 13.1 W under FD-MAC).

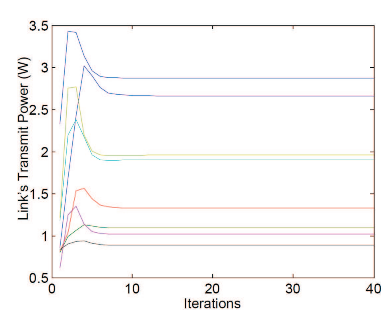
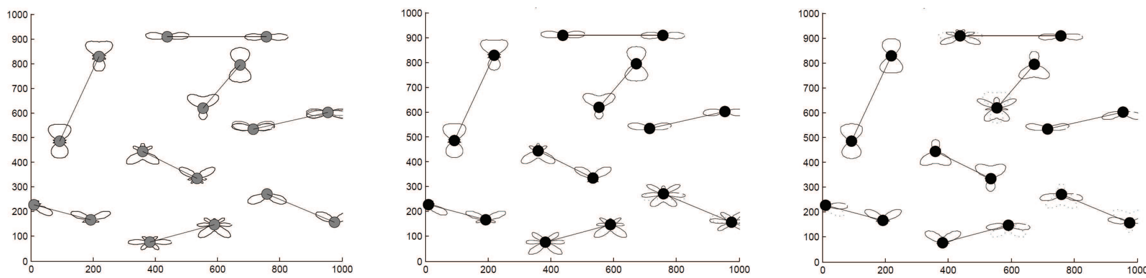


Fig. 4. Convergence of transmit power at different links under FD-MAC.



(a) Centralized Algorithm (b) FD-MAC with SPE (c) Without SPE

Fig. 2. FD radios' antenna patterns under the centralized, distributed FD-MAC (exploiting SPs), and without exploiting SPs algorithms.

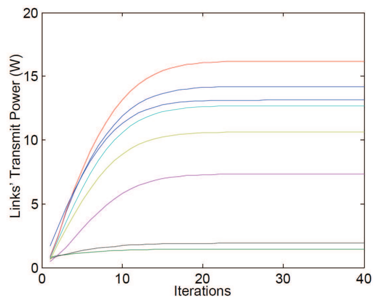


Fig. 5. Convergence of transmit power at different links when SPEs are not exploited (game (4)).

TABLE I
LINKS' TX POWER (IN WATTS).

Links	HD	CSMA-based	FD-MAC	without SPE
1	1.45	0.856	2.66	14.2
2	0.0.46	0.744	1.08	1.44
3	0.39	0.29	1.32	16.2
4	0.73	0.268	1.89	12.6
5	0.24	0.327	1.01	7.3
6	0.26	0.46	1.96	10.62
7	0.322	0.43	0.88	1.9
8	0.642	0.63	2.87	13.1

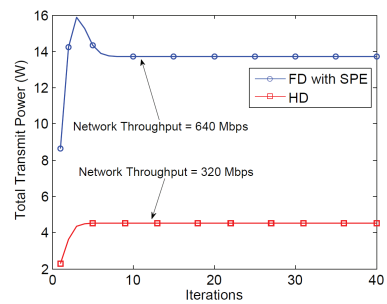


Fig. 6. Total transmit power of HD v.s. FD radios for a rate of 40 Mbps per link.

The convergence of Tx power for different links under FD-MAC is shown in Figure 4. We notice that links's Tx power under FD-MAC converge with different speeds (to the NE) but all intermediate Tx powers are in range close to Tx power under the CSMA-based approach (Table I). This is critically important as FD radios need to sustain before reaching it unique NE. If intermediate Tx power are excessive e.g., higher than nodes' power budget (like the case when SPs are not exploited, in Table I and Figure 5), nodes can not follow the game to reach its NE, even the NE can be very power efficient later. Figures V and 5 (and Table I) show that by being responsible for their interference (using (5) instead of (4)) all links can reduce their Tx power. It is also seen that radios of game (4) (in Figure 5) take long time to converge, compared with (5) (in Figure 4) as higher network interference makes them more dependent on each other and need more time to "negotiate".

Figure 6 compares the total transmit power of HD v.s. FD radios (under FD-MAC) when the rate demand is 40 Mbps for both directions of a FD link and one direction of HD link (the other direction's rate is 0). It is clear that we intend to have FD radios to produce double network throughput over HD radios but FD radios require only more than three times power than HD radios. Note that for a single FD link, double Tx power is required to double the link throughput. Moreover, as throughput is not in linearly (but asymptotically in $\log$) scale w.r.t. to Tx power, FD radios' doubling throughput at the expense of more than three time Tx power is very precious.

VI. CONCLUSIONS

By investigating the transmit power minimization problem of a FD MIMO network subject to rate demands, we established conditions under which FD radios can double the network

$$L(\mathbf{G}_i^l, \mathbf{G}_i^r, \gamma_i^l, \gamma_i^r, p) = \sum_{i=1}^{N} \{\mathrm{tr}(\mathbf{G}_i^l \mathbf{G}_i^{lH}) + \mathrm{tr}(\mathbf{G}_i^r \mathbf{G}_i^{rH}) + \frac{p}{2}((\max\{0, \gamma_i^l + pq_i^l\})^2 - (\gamma_i^l)^2 + (\max\{0, \gamma_i^r + pq_i^r\})^2 - (\gamma_i^r)^2)\} \tag{20}$$

$$0 = \frac{\partial L(\mathbf{G}_i^l, \mathbf{G}_i^r, \gamma_i^l, \gamma_i^r, p)}{\partial \mathbf{G}_j^{r*}} = 2\mathbf{G}_j^r + \frac{p}{2}\sum_{i=1}^{N}\{\frac{\partial(\max\{0, \gamma_i^l + pq_i^l\})^2}{\partial \mathbf{G}_j^{r*}} + \frac{\partial(\max\{0, \gamma_i^r + pq_i^r\})^2}{\partial \mathbf{G}_j^{r*}}\} \tag{21}$$

$$\boldsymbol{x}_i^l = \left[\Re[vec(\mathbf{G}_i^l)]^T, \Im[vec(\mathbf{G}_i^l)]^T\right]^T; \boldsymbol{x}_i^r = \left[\Re[vec(\mathbf{G}_i^r)]^T, \Im[vec(\mathbf{G}_i^r)]^T\right]^T \tag{22}$$

$$\nabla_x L = 2\left[\Re[vec(\frac{\partial L}{\partial \mathbf{G}_1^{l*}})]^T, ..., \Re[vec(\frac{\partial L}{\partial \mathbf{G}_N^{r*}})]^T, \Im[vec(\frac{\partial L}{\partial \mathbf{G}_1^{l*}})]^T, ..., \Im[vec(\frac{\partial L}{\partial \mathbf{G}_N^{r*}})]^T\right]^T \tag{23}$$

throughput over HD ones. These conditions quantify exactly the level of network interference that allows multiple FD radios/links can efficiently co-exist. Consequently, we developed a novel MAC mechanism that advocates concurrent transmissions of FD links while exploiting unique advantages of FD radios (in learning radio medium at a much finer level than just carrier sensing and the ability to instantaneously adjust/adapt transmission behavior) to reduce network interference (that, in return, facilitates FD links' co-existence). The developed MAC is fully distributed and converges to a unique NE whose efficacy is almost the same as that of the centralized algorithm.

REFERENCES

[1] J. I. Choi, M. Jain, K. Srinivasan, P. Levis, and S. Katti, "Achieving single channel, full duplex wireless communication," in *Proceedings of MobiCom Conference*, 2010, pp. 1–12.

[2] M. Jain, J. I. Choi, T. Kim, D. Bharadia, S. Seth, K. Srinivasan, P. Levis, S. Katti, and P. Sinha, "Practical, real-time, full duplex wireless," in *Proceedings of MobiCom Conference*, 2011, pp. 301–312.

[3] D. Bharadia, E. McMilin, and S. Katti, "Full duplex radios," in *Proceedings of SIGCOMM Conference*, 2013.

[4] D. Bharadia and S. Katti, "Full duplex MIMO radios," in *Proceedings of the 11th USENIX Symposium on Networked Systems Design and Implementation (NSDI)*, 2014.

[5] A. Sabharwal, P. Schniter, D. Guo, D. Bliss, S. Rangarajan, and R. Wichman, "In-band full-duplex wireless: Challenges and opportunities," *IEEE Journal on Selected Areas in Communications*, vol. 32, no. 9, pp. 1637–1652, Sept 2014.

[6] Y. Yang, B. Chen, K. Srinivasan, and N. B. Shroff, "Characterizing the achievable throughput in wireless networks with two active rf chains," in *Proceedings of the IEEE INFOCOM Conference*, 2014.

[7] X. Xie and X. Zhang, "Does full-duplex double the capacity of wireless networks," in *Proceedings of the IEEE INFOCOM Conference*, 2014.

[8] D. N. Nguyen, M. Krunz, and S. Hanly, "On the throughput of full-duplex MIMO in the multi-link case," in *Proceedings of the IEEE WiOpt Conference*, 2014.

[9] Y. Zhang, L. Lazos, K. Chen, B. Hu, and S. Shivaramaiah, "FD-MMAC: combating multi-channel hidden and exposed terminals using a single transceiver," in *Proceedings of the IEEE INFOCOM Conference*, 2014.

[10] W. Afifi and M. Krunz, "Exploiting self-interference suppression for improved spectrum awareness/efficiency in cognitive radio systems," in *Proceedings of the IEEE INFOCOM Conference*, 2013.

[11] D. Tse and P. Viswanath, *Fundamentals of Wireless Communication*. Cambridge University Press, May 2005.

[12] F. Rashid-Farrokhi, L. Tassiulas, and K. Liu, "Joint optimal power control and beamforming in wireless networks using antenna arrays," *IEEE Transactions on Communications*, vol. 46, no. 10, pp. 1313–1324, Oct 1998.

[13] R. Yates, "A framework for uplink power control in cellular radio systems," *IEEE Journal on Selected Areas in Communications*, vol. 13, no. 7, pp. 1341–1347, Sep 1995.

[14] D. Nguyen and M. Krunz, "Power minimization in mimo cognitive networks using beamforming games," *IEEE Journal on Selected Areas in Communications*, vol. 31, no. 5, pp. 916–925, May 2013.

[15] A. Auslender and M. Teboulle, *Asymptotic Cones and Functions in Optimization and Variational Inequalities*. Springer, 2003.

[16] G. Scutari, D. Palomar, F. Facchinei, and J.-S. Pang, "Convex optimization, game theory, and variational inequality theory," *IEEE Signal Processing Magazine*, pp. 35–47, May 2010.

[17] M. J. Osborne, *An Introduction to Game Theory*. Oxford University Press, 2004.

[18] A. Berman and R. Plemmons, "Nonnegative matrices in the mathematical sciences," *Society for Industrial Mathematics (SIAM)*, 1987.

[19] D. Hoang and R. Iltis, "Noncooperative eigencoding for MIMO ad hoc networks," *IEEE Trans. on Signal Processing*, vol. 56, no. 2, pp. 865–869, 2008.

[20] A. W. Marshall and I. Olkin, *Inequalities: Theory of Majorization and Its Applications*. Academic Press, 1979.

[21] R. A. Horn and C. R. Johnson, *Matrix Analysis*. Cambridge University Press, 1990.

[22] R. Iltis, S.-J. Kim, and D. Hoang, "Noncooperative iterative MMSE beamforming algorithms for ad hoc networks," *IEEE Trans. on Communications*, vol. 54, no. 4, pp. 748–759, 2006.

[23] D. Nguyen and M. Krunz, "Power-efficient spatial multiplexing for multi-antenna manets," in *Proc. of the ICC Conf.*, 2012.

[24] D. P. Bertsekas, *Nonlinear Programming*. Athena Scientific, 1995.

APPENDIX A
PROOF OF LEMMA 1

We first introduce the following proposition:

Proposition 1: Let $P^* = \text{tr}(\mathbf{G}_i^{rH}\mathbf{G}_i^r)$ and $c_i^{l*} = d_i^l$ be the transmit power by radio i_r and the throughput received by radio i_l at a NE of the game (5), then this precoder $\mathbf{G}_i^r$ is also a solution to:

$$\begin{aligned} &\underset{\{\mathbf{G}_i^r\}}{\text{maximize}} \ c_i^l \\ &\text{s.t.} \quad \text{tr}(\mathbf{G}_i^{rH}\mathbf{G}_i^r) \le P^*. \end{aligned} \tag{24}$$

Proof: If $\mathbf{G}_i^r$ is not a solution to problem (24) then there exists a precoder $\bar{\mathbf{G}}_i^r$ that requires at most power of P^* but achieves a rate $\bar{c}_i^l > d_i^l$. In other words, it is possible for radio i_r to reduce its transmit power to achieve a rate of d_i^l. This contradicts to the fact that $\mathbf{G}_i^r$ is the optimal solution of i_r at the NE where no one benefits (i.e., transmit at lower power) from unilaterally changing their strategies. Thus, $\mathbf{G}_i^r$ must be a solution of (24). □

From the definition of $\mathbb{P}_{feasible}(\boldsymbol{d})$ (6), we have the following inequality $\forall i$ radios:

$$d_i^l \ge \log|\mathbf{I} + |\mathbf{G}_i^{rH}\mathbf{H}_{ii}^{lrH}\mathbf{Q}_i^{l^{-1}}\mathbf{H}_{ii}^{lr}\mathbf{G}_i^r| \tag{25a}$$

$$\ge \log\Big(1 + \text{tr}(\mathbf{G}_i^{rH}\mathbf{G}_i^r)\text{eig}_{\max}(\mathbf{H}_{ii}^{lrH}\mathbf{Q}_i^{l^{-1}}\mathbf{H}_{ii}^{lr})\Big) \tag{25b}$$

where the (25a) comes from the Proposition 1. The RHS of (25b) is a lower-bound of the rate c_i^l in problem (24), obtained by allocating all power $\text{tr}(\mathbf{G}_i^{rH}\mathbf{G}_i^r)$ on the subchannel $\text{eig}_{\max}(\mathbf{H}_{ii}^{lrH}\mathbf{Q}_i^{l^{-1}}\mathbf{H}_{ii}^{lr})$ and zero power on others.

Besides, let $\mathbf{V}_{ii}^l$ be the unitary matrix that diagonalizes matrix $\mathbf{H}_{ii}^{lrH}\mathbf{Q}_i^{l^{-1}}\mathbf{H}_{ii}^{lr}$ and the diagonal matrix $\mathbf{W}_i^l$ contain eigenvalues of matrix $\mathbf{H}_{ii}^{lrH}\mathbf{Q}_i^{l^{-1}}\mathbf{H}_{ii}^{lr}$, we have:

$$\mathbf{V}_{ii}^{lH}\mathbf{H}_{ii}^{lrH}\mathbf{Q}_i^{l^{-1}}\mathbf{H}_{ii}^{lr}\mathbf{V}_{ii}^l = \mathbf{W}_i^l \tag{26a}$$

$$\text{tr}(\mathbf{Q}_i^l) = \text{tr}(\mathbf{H}_{ii}^{lr}\mathbf{V}_{ii}^l\mathbf{W}_i^{l-1}\mathbf{V}_{ii}^{lH}\mathbf{H}_{ii}^{lrH}) \tag{26b}$$

Then we have equations (27b),(27c),(27d),(27e). (27b) follows by recalling the noise-plus-interference covariance matrix on the LHS of (26b) and applying the identity $\frac{\text{tr}(\mathbf{A})}{n} \ge |\mathbf{A}|^{1/n}$ [21] (for any $n \times n$ positive semi-definite matrix $\mathbf{A}$) to the RHS of (26b). (27d) follows from applying the identity $\text{tr}(\mathbf{AB}) \le \text{tr}(\mathbf{A})\text{tr}(\mathbf{B})$ to the LHS of (27c).

From inequalities (27e) and (25b), we get (27a). We then have:

$$\begin{aligned} &(2^{d_i^l} - 1) \ge \boldsymbol{\Gamma}(2i-1,:)\times \\ &\quad [\text{tr}(\mathbf{G}_1^{lH}\mathbf{G}_1^l), \text{tr}(\mathbf{G}_1^{rH}\mathbf{G}_1^r), \ldots, \text{tr}(\mathbf{G}_N^{lH}\mathbf{G}_N^l), \text{tr}(\mathbf{G}_N^{rH}\mathbf{G}_N^r)]^T \\ &\boldsymbol{\Gamma}^{-1}\times[2^{d_1^l}-1, 2^{d_1^r}-1, \ldots, 2^{d_N^l}-1, 2^{d_N^r}-1]^T \ge \\ &\quad [\text{tr}(\mathbf{G}_1^{lH}\mathbf{G}_1^l), \text{tr}(\mathbf{G}_1^{rH}\mathbf{G}_1^r), \ldots, \text{tr}(\mathbf{G}_N^{lH}\mathbf{G}_N^l), \text{tr}(\mathbf{G}_N^{rH}\mathbf{G}_N^r)]^T. \end{aligned} \tag{28a}$$

where the inequality (28a) comes from the assumption $\boldsymbol{\Gamma}$ is a P-matrix, hence invertible [18].

Hence the RHS in (28a) is bounded or rate demands can be fullfilled with a bounded power allocation vector $\boldsymbol{p}$. In other words, $\mathbb{P}_{feasible}(\boldsymbol{d})$ contains at least one bounded $\boldsymbol{p}$. □

APPENDIX B
PROOF OF LEMMA 2

For $\boldsymbol{f} \in \mathbb{Q}^{asymp}(\boldsymbol{d})$, by the definition of limit directions, sequences $\{\boldsymbol{p_n}\}$ and $\{\nu_n\}$ exist so that we have (27f). Then, (27g), (27h), and (27i) follow from (27e), (10), and the definition of $\boldsymbol{d}$, respectively. Lemma 2 is proved. □

$$d_i^l \geq \log(1+ \frac{\text{tr}(\mathbf{G}_i^{rH}\mathbf{G}_i^r)|\mathbf{H}_{ii}^{lrH}\mathbf{H}_{ii}^{lr}|^{\frac{1}{M}}}{1+ g_{sis}\frac{\text{tr}(\mathbf{H}_{ii}^{llH}\mathbf{H}_{ii}^{ll})}{M}\text{tr}(\mathbf{G}_i^{lH}\mathbf{G}_i^l) + \sum_{j=1|j\neq i}^{N}(\frac{\text{tr}(\mathbf{H}_{ij}^{lrH}\mathbf{H}_{ij}^{lr})}{M}\text{tr}(\mathbf{G}_j^{rH}\mathbf{G}_j^r) + \frac{\text{tr}(\mathbf{H}_{ij}^{llH}\mathbf{H}_{ij}^{ll})}{M}\text{tr}(\mathbf{G}_j^{lH}\mathbf{G}_j^l))}) \tag{27a}$$

$$M + g_{sis}\text{tr}(\mathbf{H}_{ii}^{ll}{\mathbf{H}_{ii}^{ll}}^H\mathbf{G}_i^l{\mathbf{G}_i^l}^H) + \sum_{j=1|j\neq i}^{N}(\text{tr}(\mathbf{H}_{ij}^{llH}\mathbf{H}_{ij}^{ll}\mathbf{G}_j^{lH}\mathbf{G}_j^l) + \text{tr}(\mathbf{H}_{ij}^{lrH}\mathbf{H}_{ij}^{lr}\mathbf{G}_j^{rH}\mathbf{G}_j^r)) \geq M|\mathbf{H}_{ii}^{lr}\mathbf{V}_i^l{\mathbf{W}_i^l}^{-1}\mathbf{V}_i^{lH}\mathbf{H}_{ii}^{lrH}|^{\frac{1}{M}} \tag{27b}$$

$$M + g_{sis}\text{tr}(\mathbf{H}_{ii}^{ll}{\mathbf{H}_{ii}^{ll}}^H\mathbf{G}_i^l{\mathbf{G}_i^l}^H) + \sum_{j=1|j\neq i}^{N}(\text{tr}(\mathbf{H}_{ij}^{llH}\mathbf{H}_{ij}^{ll}\mathbf{G}_j^{lH}\mathbf{G}_j^l) + \text{tr}(\mathbf{H}_{ij}^{lrH}\mathbf{H}_{ij}^{lr}\mathbf{G}_j^{rH}\mathbf{G}_j^r)) \geq M|\mathbf{H}_{ii}^{lrH}\mathbf{H}_{ii}^{lr}|^{\frac{1}{M}}\frac{1}{\text{eig}_{\max}(\mathbf{W}_i^l)} \tag{27c}$$

$$M + g_{sis}\text{tr}(\mathbf{H}_{ii}^{ll}{\mathbf{H}_{ii}^{ll}}^H)\text{tr}(\mathbf{G}_i^l{\mathbf{G}_i^l}^H) + \sum_{j=1|j\neq i}^{N}(\text{tr}(\mathbf{H}_{ij}^{llH}\mathbf{H}_{ij}^{ll})\text{tr}(\mathbf{G}_j^{lH}\mathbf{G}_j^l) + \text{tr}(\mathbf{H}_{ij}^{lrH}\mathbf{H}_{ij}^{lr})\text{tr}(\mathbf{G}_j^{rH}\mathbf{G}_j^r)) \geq M|\mathbf{H}_{ii}^{lrH}\mathbf{H}_{ii}^{lr}|^{\frac{1}{M}}\frac{1}{\text{eig}_{\max}(\mathbf{W}_i^l)} \tag{27d}$$

$$\text{eig}_{\max}(\mathbf{W}_i^l) \geq \frac{|\mathbf{H}_{ii}^{lrH}\mathbf{H}_{ii}^{lr}|^{\frac{1}{M}}}{1 + g_{sis}\frac{\text{tr}(\mathbf{H}_{ii}^{llH}\mathbf{H}_{ii}^{ll})}{M}\text{tr}(\mathbf{G}_i^{lH}\mathbf{G}_i^l) + \sum_{j=1|j\neq i}^{N}(\frac{\text{tr}(\mathbf{H}_{ij}^{llH}\mathbf{H}_{ij}^{ll})}{M}\text{tr}(\mathbf{G}_j^{lH}\mathbf{G}_j^l) + \frac{\text{tr}(\mathbf{H}_{ij}^{lrH}\mathbf{H}_{ij}^{lr})}{M}\text{tr}(\mathbf{G}_j^{rH}\mathbf{G}_j^r))} \tag{27e}$$

$$d_i^l = c_i^l(\boldsymbol{p}_n) \geq \log\left(1 + \text{tr}(\mathbf{G}_i^{rH}\mathbf{G}_i^r)\text{eig}_{\max}(\mathbf{H}_{ii}^{lrH}{\mathbf{Q}_i^l}^{-1}\mathbf{H}_{ii}^{lr})\right) \tag{27f}$$

$$\geq \log(1+ \frac{\text{tr}(\mathbf{G}_i^{rH}\mathbf{G}_i^r)|\mathbf{H}_{ii}^{lrH}\mathbf{H}_{ii}^{lr}|^{\frac{1}{M}}}{1 + g_{sis}\text{tr}(\mathbf{H}_{ii}^{ll}{\mathbf{H}_{ii}^{ll}}^H)\text{tr}(\mathbf{G}_i^l{\mathbf{G}_i^l}^H) + \sum_{j=1|j\neq i}^{N}(\frac{\text{tr}(\mathbf{H}_{ij}^{llH}\mathbf{H}_{ij}^{ll})}{M}\text{tr}(\mathbf{G}_j^{lH}\mathbf{G}_j^l) + \frac{\text{tr}(\mathbf{H}_{ij}^{lrH}\mathbf{H}_{ij}^{lr})}{M}\text{tr}(\mathbf{G}_j^{rH}\mathbf{G}_j^r))}) \tag{27g}$$

$$= \log(1+ \frac{\text{tr}(\mathbf{G}_i^{rH}\frac{1}{\nu_n}\mathbf{G}_i^r)|\mathbf{H}_{ii}^{lrH}\mathbf{H}_{ii}^{lr}|^{\frac{1}{M}}}{\frac{1}{\nu_n} + g_{sis}\text{tr}(\mathbf{H}_{ii}^{ll}{\mathbf{H}_{ii}^{ll}}^H)\text{tr}(\mathbf{G}_i^l\frac{1}{\nu_n}{\mathbf{G}_i^l}^H) + \sum_{j=1|j\neq i}^{N}(\frac{\text{tr}(\mathbf{H}_{ij}^{llH}\mathbf{H}_{ij}^{ll})}{M}\text{tr}(\mathbf{G}_j^{lH}\frac{1}{\nu_n}\mathbf{G}_j^l) + \frac{\text{tr}(\mathbf{H}_{ij}^{lrH}\mathbf{H}_{ij}^{lr})}{M}\text{tr}(\mathbf{G}_j^{rH}\frac{1}{\nu_n}\mathbf{G}_j^r))}) \tag{27h}$$

$$\overset{n\to+\infty}{=} {c'}_i^l(\boldsymbol{f}) \qquad \forall i_l, i_r. \tag{27i}$$

APPENDIX C
PROOF OF THEOREM 3

Variational Inequality (VI) problem: [16] Given a subset $\mathbb{K}$ of the Euclidean n-dimensional space $\mathbb{R}^n$ and a mapping $F : \mathbb{K} \to \mathbb{R}^n$, a VI$(\mathbb{K}, \mathbb{R}^n)$ problem is to find a vector $\boldsymbol{x^{opt}} \in \mathbb{K}$ so that:

$$(\boldsymbol{x} - \boldsymbol{x^{opt}})^T F(\boldsymbol{x^{opt}}) \geq 0, \quad \forall \boldsymbol{x} \in \mathbb{K}. \tag{29}$$

If the set $\mathbb{K}$ has a Cartesian structure, i.e., $\mathbb{K} = \mathbb{K}_1 \times \mathbb{K}_2 \times \ldots \times \mathbb{K}_N$ (where $\mathbb{K}_i \in \mathbb{R}^{n_i}$ and $\sum_{i=1}^{N} n_i = n$), we have the following theorem regarding the existence and uniqueness of a solution to the above VI problem [16].

Theorem 5: If set $\mathbb{K}$ has a Cartesian structure, the VI$(\mathbb{K}, \mathbb{R}^n)$ problem has a unique solution $\boldsymbol{x}^{opt}$ provided $\mathbb{K}_u$ is closed and convex and F is continuous uniformly-P function, i.e, there exists a positive constant α such that:

$$\max_{\{1\leq i\leq N\}} (\boldsymbol{x}_i - \boldsymbol{x}'_i)^T(F(\boldsymbol{x}_i) - F(\boldsymbol{x}'_i)) \geq \alpha\|\boldsymbol{x}_i - \boldsymbol{x}'_i\|^2, \quad \forall \boldsymbol{x}_i, \boldsymbol{x}'_i \in \mathbb{K}_i. \tag{30}$$

To cast the game (5) as a VI problem, we use the $vec()$ operator in (22) to map the complex matrix in (5) to the Euclidean domain, by stacking columns (from left to right) of an $m \times n$ matrix to form an $mn \times 1$ vector. The gradient of a matrix function (.) w.r.t $\mathbf{G}_i^r$ is in (23).

If the condition in Theorem 2 holds, the strategic space of player i_r, denoted by $Q_i^r \in \mathbb{C}^{M\times M}$, is nonempty. Moreover, it can be verified that Q_i^r is convex and bounded. Hence, problem (5) is convex. The following inequality captures the necessary (and also the sufficient) condition for strategy $\hat{\mathbf{G}}_i^r$ to be the optimal response:

$$(\mathbf{G}_i^r - \hat{\mathbf{G}}_i^r) \bullet \nabla U_i^r \geq 0 \quad \forall \mathbf{G}_i^r \in Q_i \tag{31}$$

where $\mathbf{A} \bullet \mathbf{B} \overset{\text{def}}{=} \text{vec}(\mathbf{A})^T\text{vec}(\mathbf{B})$ and $U_i^r \overset{\text{def}}{=} \text{tr}(\mathbf{G}_i^r{\mathbf{Q}_i^r}^T\mathbf{G}_i^{rH})$

Define $Q \overset{\text{def}}{=} Q_1^l \times Q_1^r \ldots Q_N^l \times Q_N^r$ and $F \overset{\text{def}}{=} F_1^l \times F_1^r \ldots \times F_N^l \times F_N^r$ with $F_i^l \overset{\text{def}}{=} \nabla U_i^l$. By comparing (31) with the definition of a VI problem, the set $\hat{\mathbf{G}} \overset{\text{def}}{=} [\hat{\mathbf{G}}_1^l \times \hat{\mathbf{G}}_1^r \ldots \times \hat{\mathbf{G}}_N^l \times \hat{\mathbf{G}}_N^r]$ is a NE of the game (5) iff $\hat{\mathbf{G}}$ is a solution of the VI(Q, F) problem.

Let $\mathbf{G} \overset{\text{def}}{=} [\mathbf{G}_1^l \times \mathbf{G}_1^r \ldots \times \mathbf{G}_N^l \times \mathbf{G}_N^r]$ and $\hat{\mathbf{G}} \overset{\text{def}}{=} [\hat{\mathbf{G}}_1^l \times \hat{\mathbf{G}}_1^r \ldots \times \hat{\mathbf{G}}_N^l \times \hat{\mathbf{G}}_N^r]$ be two different strategy sets of the strategic space Q of the game (5), then:

$$F(\hat{\mathbf{U}}_i^r) = [(\mathbf{S}_i^r + \mathbf{I})\hat{\mathbf{G}}_i^r] \tag{32}$$

Consequently:

$$\begin{aligned} &\text{vec}(\hat{\mathbf{G}}_i^r - \mathbf{G}_i^r)^T\text{vec}(F(\hat{\mathbf{G}}_i^r) - F(\mathbf{G}_i^r)) \\ &\geq \text{eig}_{\min}(\mathbf{S}_i^r + \mathbf{I})||\text{vec}((\hat{\mathbf{G}}_i^r - \mathbf{G}_i^r))||^2 \end{aligned} \tag{33}$$

where (33) follows from the fact that $||\mathbf{A}\boldsymbol{a}|| \geq \text{eig}_{\min}(\mathbf{A})||\boldsymbol{a}||$.

Since $\mathbf{S}_i^r + \mathbf{I}$ is positive definite, $\text{eig}_{\min}(\mathbf{S}_i^r + \mathbf{I}) > 0$. Hence, VI$(Q, \bar{F})$ problem or game (5) has a unique NE. $\square$